Molecular Characterization of Maize Inbreds with Expired U.S. Plant Variety Protection

Paul T. Nelson, Nathan D. Coles, James B. Holland, David M. Bubeck, Stephen Smith, and Major M. Goodman*

ABSTRACT

Maize inbred lines with expired Plant Variety Protection Act (PVPA) certificates are publicly available and potentially represent a new germplasm resource for many public and private breeding programs. However, accurate pedigree and genetic background information for ex-PVPA maize inbreds is necessary if they are to be effectively utilized in breeding efforts. We have used single nucleotide polymorphism (SNP) markers to evaluate the relationships and population structure among 92 ex-PVPA inbred lines in relation to 17 well-known public inbreds. Based on unweighted pair group method with arithmetic mean clustering, principal components analysis, and model-based clustering, we identified six primary genetic clusters represented by the prominent inbred lines B73, Mo17, PH207, A632, Oh43, and B37. We also determined the genetic background of ex-PVPA inbreds with conflicting, ambiguous, or undisclosed pedigrees. We assessed genetic diversity across subsets of ex-PVPA lines and concluded that the ex-PVPA lines are no more diverse than the public set evaluated here. Additionally, all alleles present in the ex-PVPA inbreds, for the 614 SNPs included in this study, are also found in public temperate maize germplasm.

P.T. Nelson, N.D. Coles, and M.M. Goodman, Dep. of Crop Science, North Carolina State Univ., Box 7620, Raleigh, NC 27695; J.B. Holland, USDA-ARS, Plant Science Research Unit, Dep. of Crop Science, North Carolina State Univ., Raleigh, NC 27695-7620; D.M. Bubeck and S. Smith, Pioneer Hi-Bred International, Inc., Crop Genetics Research and Product Development, DuPont Agriculture and Nutrition, 7300 NW 62nd Ave., Box 1004, Johnston, IA 50131-1004. Received 13 Feb. 2008. *Corresponding author (maize_resources@ncsu.edu).

Abbreviations: DK, DEKALB Genetics; GEM, Germplasm Enhancement of Maize; LH, Holden's Foundation Seeds; NK, Novartis Seeds; PCA, principal components analysis; PH, Pioneer Hi-Bred International Inc.; PVPA, Plant Variety Protection Act; SNP, single nucleotide polymorphism; UPGMA, unweighted pair group method with arithmetic mean.

In the past three decades, maize breeding in the United States and abroad has become increasingly proprietary in nature. Since the mid-1970s, commercial investment in improving performance of maize hybrids has increased at least 10-fold, possibly more when biotechnology and transgenic trait development activities are considered. The total commercial global investment in improvement of maize on a commercial scale is over one billion dollars annually. Whereas inbreds from public institutions formerly played a major role in private breeding and seed production, modern public inbreds currently contribute very little. This is illustrated in a series of surveys conducted by the American Seed Trade Association from 1956 to 1986 (Sprague, 1971; Zuber, 1975; Zuber and Darrah, 1980; Darrah and Zuber, 1986). In the 1970 survey, 71.9% of commercial hybrid seed production included at least one publicly developed parent; in 1985, the year of the most recent survey, this number had dropped to 37.9%.

Published in Crop Sci. 48:1673–1685 (2008).
doi: 10.2135/cropsci2008.02.0092

677 S. Segoe Rd., Madison, WI 53711 USA

Today, such surveys are effectively impossible to conduct due to trade secrets and intellectual property restrictions, but it is likely that <1% of all maize hybrids sold in the United States contain a publicly developed inbred. This shift has been accompanied by a decrease in the number of public maize breeding programs that are actively releasing lines (Coors, 2006). The germplasm pool available to public programs is shrinking in terms of current germplasm, and public programs are arguably two decades or more behind private breeding programs in performance for yield and other traits of agronomic importance. Additionally, public inbreds often lack inbred per se performance for traits of commercial importance such as female yield, seed quality, and emergence.

As competition in the hybrid maize seed industry has intensified, the control of germplasm has become increasingly important. In 1970, the U.S. Plant Variety Protection Act (PVPA; 7 U.S.C. §§ 2321–2583) was passed. It states (p. 14) that "the breeder (or the successor in interest of the breeder), has the right, during the term of the plant variety protection, to exclude others from selling the variety, or offering it for sale, or reproducing it, or importing it, or exporting it, or using it in producing (as distinguished from developing) a hybrid or different variety therefrom, to the extent provided by this Act." For maize inbreds, the PVPA initially provided 18 yr of protection but in 1994 protection was extended to 20 yr from the date that the certificate was issued; on expiration, lines are made available to the public provided that any applicable U.S. utility patent protection has expired. From 2001 through 2007, there were at least 114 maize PVPA certificate expirations, with accumulation of more releases each additional year. This represents a large and growing germplasm resource, previously unavailable to many breeding programs, that can be utilized in breeding within the public and private sectors.

Effective utilization of maize germplasm in breeding requires accurate characterization of line performance and line relationships to other germplasm. When developing breeding populations, maize breeders should choose parents that (i) show superior performance for the traits of interest, (ii) maximize within-population variance for the traits of interest, and (iii) preserve heterotic patterns for maximum heterosis in hybrid development. To do this, breeders require accurate phenotypic data on potential parents and an understanding of the relationships among these lines. Mikel and Dudley (2006) and Mikel (2006) reviewed inbreds under PVPA and U.S. Patent protection during the period from 1980 to 2006, including inbreds with expired PVPA certificates. Using pedigree information, primarily from PVPA and U.S. Patent records, they inferred derivation and relative usage of these inbreds. This information is of great value and marks a logical starting point in the characterization of PVPA germplasm. However, the utility of pedigree records is limited, especially when making inferences among lines that span multiple breeding programs, have been subjected to selection, or are derived from recurrent selection populations (Bernardo, 1993). Furthermore, PVPA certificate pedigrees are often vague or refer to progenitors of undisclosed (and occasionally erroneous) origin. Molecular markers can partially compensate for such inadequacies in pedigree records, providing an additional level of accuracy in estimating relationships and population structure.

Our objective is to build on previous work by Mikel (2006) by evaluating population structure, line derivation, and relationships among newly available ex-PVPA maize inbreds using molecular markers. We also assess genetic diversity among the ex-PVPA inbreds in relation to key progenitor public inbreds. The results of this characterization will benefit maize breeders from public and private organizations by assisting in the effective use of this newly available germplasm resource.

MATERIALS AND METHODS

Germplasm

A set of 92 ex-PVPA maize inbreds was chosen for genotyping. All of the lines described by Mikel (2006) were included in this study except seven: DKMBPM, LH143, LP1CmsHt, PHK05, PHK29, PHR25, and PHV78. Ten ex-PVPA lines that were not described by Mikel (2006) were included; these are PVPA lines with more recently expired certificates or that have abandoned or withdrawn PVPA status (Table 1). Copies of the PVPA certificates are available online from the Plant Variety Protection Office (http://www.ams.usda.gov/pvpo, verified 8 Apr. 2008). Descriptive information for NK792, the single line for which the PVPA application was withdrawn before certificate issuance, is not publicly available. Seed for the ex-PVPA lines was obtained from Mark Millard at the North Central Regional Plant Introduction Station (Ames, IA) and is available on request through the Germplasm Resource Information Network (http://www.ars-grin.gov/npgs, verified 8 Apr. 2008). In addition to ex-PVPA lines, 17 public inbreds were genotyped. These inbreds represent a broad cross-section of public U.S. maize germplasm, and many are progenitors of the ex-PVPA lines. Seed for public inbreds was from various sources (Table 2). Inbred lines B73, Mo17, and the hybrid B73 × Mo17 were included in replicate (4, 4, and 5 samples, respectively) to serve as standards, both in genotyping and in analysis. In total, there were 109 inbreds included in the study and, including B73 × Mo17 and replicate samples, there were 120 samples genotyped. All 120 samples and seed sources are listed in Supplementary Table S1.

Genotyping

Genotyping was performed using the Illumina GoldenGate high-throughput assay (Fan et al., 2003) on 768 bi-allelic single nucleotide polymorphism (SNP) markers. The markers were made available by Pioneer Hi-Bred International Inc. (Johnston, IA). Seventy-five of the markers failed outright and are not given further mention in this paper. The markers used in genotyping were chosen because they had expected

Table 1. Maize inbred lines genotyped in this study that were not described by Mikel (2006). Except where otherwise indicated, information was gathered from Plant Variety Protection Act (PVPA) certificates.

Line	PVPA certificate no.	Applicant†	Background‡	50% silking date (HU§)	Kernel row no.	Endosperm color	Cob color	Derivation	PVPA status
11430	008800177	Cargill	Lan	1335	16	Yellow	White	OH43, H99, Mo17	Expired
DK2MA22	008800193	DEKALB	Lan	1510	14	Yellow	Pink	4780 × 5P9-1	Expired
DKMBST	008800194	DEKALB	Lan	1514	12	Yellow	White	LH38 × 4726-1	Expired
LH150	008500153	Holden's	UR	1711	18	Yellow	Red	Pioneer 3147	Abandoned
ML606	009400242	United AgriSeeds	Lan	1519	13	Yellow	White	LK2 × LH38	Abandoned
NK792	NSL 243016¶	Novartis	UK	1611#	12#	Yellow#	Red#	NA††	Withdrawn
NKW8304	008800032	Novartis	SS	1600	18	Yellow	Red	B14A^2 × B73	Expired
NQ508	009200061	United AgriSeeds	UR	1421	12	Yellow	White	P3713	Abandoned
NS501	008800149	United AgriSeeds	SS	1764#	12	White	Red	A634 × K1	Expired
PHH93	008800216	Pioneer	IO	1430	16	Yellow	Red‡‡	PH806 × PH207	Expired

†DEKALB, DEKALB Genetics; Holden's, Holden's Foundation Seeds; Novartis, Novartis Seeds; Pioneer, Pioneer Hi-Bred Int.

‡Background: Stiff Stalk (SS), Lancaster (Lan), Iodent (IO), unrelated (UR), unknown (UK).

§Method for heat unit (HU) calculation varies by certificate. Where observed in North Carolina, heat units were calculated as growing degree-days as given by McMaster and Wilhelm (1997).

¶National Seed Laboratory (NSL) Source, application withdrawn before PVPA certificate was issued.

#Based on observation in North Carolina.

††Information not available.

‡‡The PVPA certificate for PHH93 indicates brown cobs, but cobs are red in observations in North Carolina.

heterozygosity values >0.2 for at least one germplasm group in a previous germplasm survey of a number of public and Pioneer germplasm groups including the following: European flints, selected tropical lines from Latin America, U.S. non–Stiff Stalk lines, and U.S. Stiff Stalk lines. Therefore, the SNP should be a reasonably unbiased set of maize markers to use across different germplasm groups and this limited set of ex-PVPA lines. Marker names, including positions (where known), are listed in Supplementary Table S2 and additional information for many of the markers is available at the Panzea website (http://www.panzea.org, verified 8 Apr. 2008).

Six of the public inbreds were genotyped by the Panzea project (Table 2) with a set of SNPs that has 602 markers in common with our set of 693, with the exception of B14, for which only 370 markers are in common. Data for these six were retrieved from the Panzea website and genotypic values were set to "missing" for the 91 deficient markers (323 deficient markers for B14).

PowerMarker (Liu and Muse, 2005) was used to eliminate markers with >25% missing data or >10% heterozygous allele calls. In total, 614 markers were used for analysis, 544 of which were in common between our data set and the data set from Panzea. Genotypic data for the 693 SNP on 115 inbreds are available in Supplementary Table S3 (not including data that originated from Panzea).

Data Analysis

A dendrogram (Fig. 1) was constructed using NTSYSpc 2.2 (Rohlf, 2007). Clustering was performed using unweighted pair group method with arithmetic mean (UPGMA) from a Jaccard's similarity coefficient matrix. The Jaccard's similarity matrix for all 120 samples is available in Supplementary Table S4.

Principal components analysis (PCA; Fig. 2 and 3) was performed using PROC PRINCOMP in SAS version 9.1.3 (SAS Institute, 2003). For the PRINCOMP procedure, bi-allelic nucleotide calls were coded 1 or –1 and missing data were assigned 0 values for this analysis.

The program Structure 2.1 (Pritchard et al., 2000) was used to assign lines to populations of similar genetic structure. Structure 2.1 employs a model-based clustering method that accounts for Hardy–Weinberg and linkage disequilibrium and attempts to find population groupings that minimize these disequilibria (Pritchard et al., 2000). Duplicate lines and B73 × Mo17 were excluded from this analysis. Structure 2.1 was

Table 2. Maize public lines included in genotyping, listed with seed sources.

Line	Seed source	Line	Seed source
A632	Panzea†	C103	NCSU
B14	Panzea	Gaspé Flint#	NCSU
B37	NCSU‡	GT112	NCSU
B52	Panzea	H99	Panzea
B73	NCSU	Mo17	NCSU
B73-P1§	Pioneer¶	Mo17-P1	Pioneer
B73-P2	Pioneer	Mo17-P2	Pioneer
B73-P3	Pioneer	Mo17-P3	Pioneer
B73 × Mo17-1	Pioneer	NC258	NCSU
B73 × Mo17-2	Pioneer	NC268	NCSU
B73 × Mo17-3	Pioneer	Oh43	Panzea
B73 × Mo17-4	Pioneer	Pa91	NCSU
B73 × Mo17-5	Pioneer	SC76	NCSU
B84	Panzea	Va35	NCSU

†Genotypic data were retrieved from the Panzea website (http://www.panzea.org, verified 8 Apr. 2008). Readers are referred to the Panzea website for seed source information because genotyping was often done on multiple seed lots.

‡North Carolina State University.

§A hyphen, "-", indicates a replication of the base line, i.e., B73-P3 is the third replicate of B73 with a Pioneer Hi-Bred seed source.

¶Pioneer Hi-Bred Int.

#Gaspé Flint is an open pollinated variety.

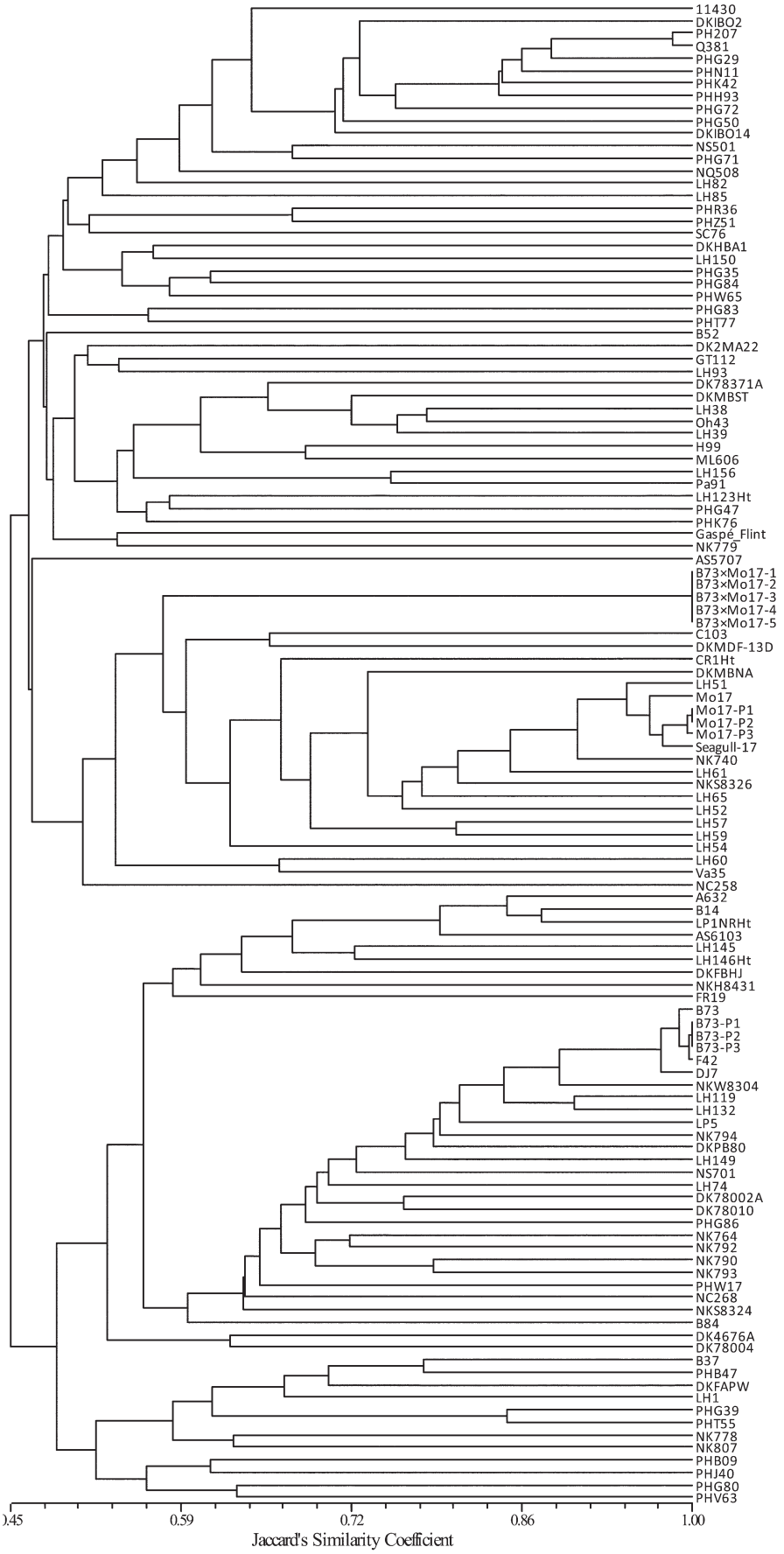

Figure 1. Dendrogram from UPGMA clustering for 92 ex–Plant Variety Protection Act maize lines, 17 public inbreds, and B73 × Mo17, using Jaccard's similarity coefficient.

run with population number, K, ranging from 1 to 11, and was repeated five times for each value of K. For each run, the number of burn-in and replication cycles was set to 500,000 each. Maximum likelihood estimates were used to select a model with an optimal population number. Lines were assigned to populations by averaging admixture values, Q, across the five repeated runs at the optimal K value (Supplementary Table S5). Q_{ik} represents the estimated proportion of the genome of line i derived from population k (Pritchard et al., 2000). A bar plot of line membership in each population was generated using Structure 2.1 (Fig. 4).

Genetic diversity was quantified via average gene diversity per locus, average Jaccard's distance between lines, average allele number per locus, number of unique alleles among subsets of lines, and number of absent alleles among subsets of lines. Gene diversity is defined for the r^{th} locus as $D_r = 1 - \sum_{i=1}^{m} x_i^2$, where m is the number of alleles and x_i is the population frequency of the ith allele at locus r (Nei, 1987). Average Jaccard's distance, $\bar{D}_J$, was calculated as $\bar{D}_J = 1 - \bar{J}$, where $\bar{J}$ is the average Jaccard's similarity coefficient. Allelic diversity among subsets was quantified by measuring (i) average number of alleles per locus; (ii) unique alleles, alleles that are present in only one of the six subsets; (iii) absent alleles, alleles that are present in all but one of the subsets; and (iv) total absent alleles, the total number of alleles that are absent in a given subset but present in at least one other subset. Because estimates of average allele number, unique alleles, absent alleles, and total absent alleles are influenced by the number of lines in the subset, a resampling technique was used to eliminate the effect of sample size. Five thousand random samples (Lu and Bernardo, 2001) of 13 lines, which is the number of lines in the smallest subset, were drawn without replacement from all subsets with >13 lines. Average allele number, unique alleles, absent alleles, and total absent alleles were calculated for each of the 5000 samples and then averaged across samples.

Genetic diversity was assessed for the entire set of ex-PVPA lines. Lines were also subdivided by applicant company (as opposed to company of current ownership, noting more recent consolidation in the hybrid seed corn industry). Where applicant companies had fewer than five ex-PVPA lines, their lines were pooled into a single subset of "other" companies. Public inbreds were combined into a single "public" subset. Duplicate entries and B73 × Mo17 were excluded from

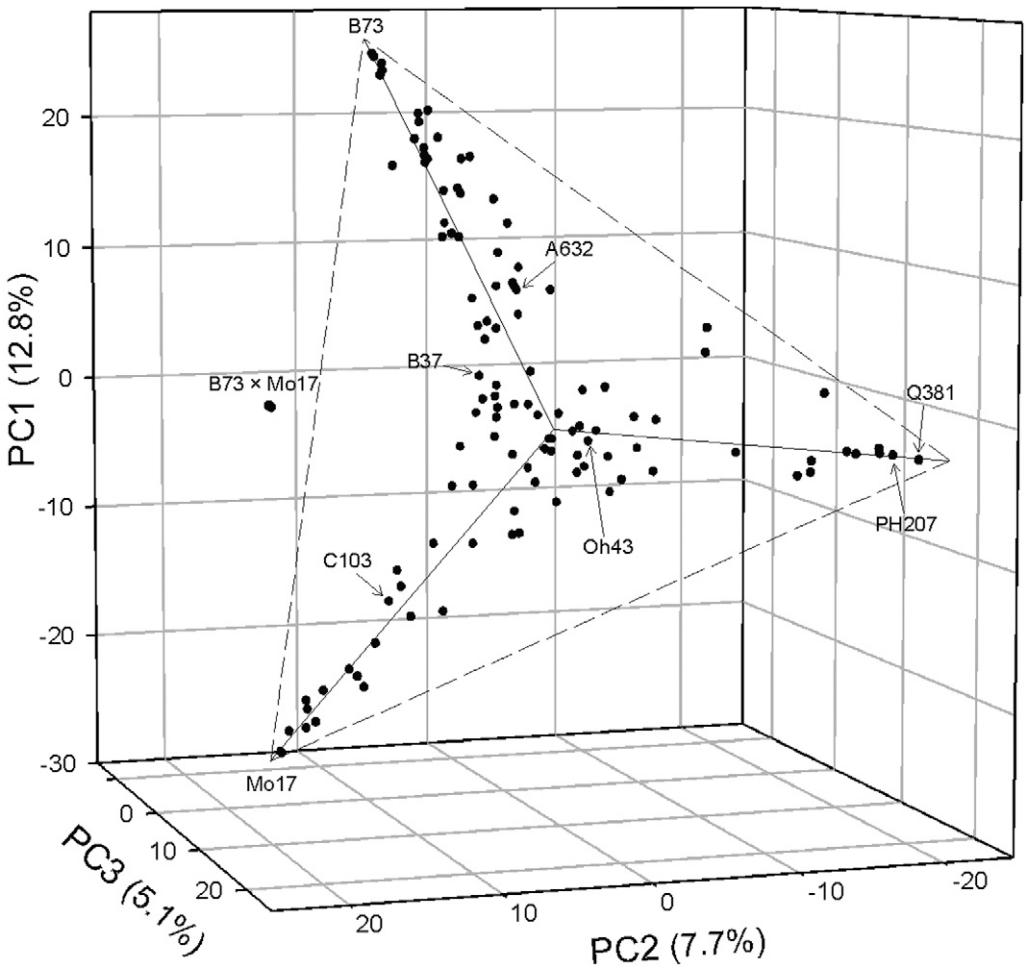

Figure 2. Scatter plot of the first three principal components (PC1, PC2, PC3) from principal components analysis for ex–Plant Variety Protection Act and public maize lines. B73 (Stiff Stalk), Mo17 (Lancaster), and PH207 (Iodent) clusters form the edges of a tetrahedron-shaped cloud. The first three principal components explain 25.5% of the total variation across 614 SNPs.

diversity analyses. Gaspé Flint, GT112, NC258, and SC76 were also excluded from diversity analysis because these have not been used extensively in midwestern U.S. maize breeding and may contribute to genetic diversity in a way that makes comparison with the other subsets less meaningful. Thus, six subsets were used in diversity comparisons, represented by DEKALB (DK; DEKALB Genetics, DeKalb, IL), Holden's (LH; Holden's Foundation Seeds, Williamsburg, IA), Pioneer (PH; Pioneer Hi-Bred International Inc.), Novartis (NK; Novartis Seeds, Golden Valley, MN), other, and public (Table 3). Where applicable, multiple comparisons for diversity analysis were made using a Tukey adjustment (Steel et al., 1997, p. 191).

For marker loci at which the ex-PVPA lines contain alleles that are absent in the public subset, we also expanded the sample of public lines to include all public temperate inbreds in the Panzea database for which allelic information was available. Depending on the marker being studied, the Panzea database contained genotypic data for 50 to 213 public temperate inbred lines.

RESULTS AND DISCUSSION

Cluster Analysis

A dendrogram from UPGMA cluster analysis reveals six predominant clusters that are represented by the following lines: B73, Mo17, PH207 (Iodent), A632, Oh43, and B37 (Fig. 1). Several lines of mixed or unrelated background, as well as B73 × Mo17, fail to cluster with any one of the aforementioned groups. These results generally agree with conclusions by Mikel and Dudley (2006), who stated that much of today's germplasm originates from seven progenitor lines: B73, Mo17, PH207, PHG39 (a B37 derivative), LH123Ht, LH82, and PH595. In our analysis, LH123Ht, which was selfed out of a hybrid, Pioneer 3535, clusters with the Oh43-type lines. Over the past 25 yr, several prominent maize breeders have drawn similar conclusions about influential maize inbreds. Inbreds cited as prominent in seven papers are listed in Table 4, where B73 and Mo17 are the most referenced lines, followed by A632, B37, and Oh43. While the Iodents are clearly an important component of modern maize production, they were not mentioned by most of the authors because the modern Iodents were originally exclusive to Pioneer and, until recently, were little known among public breeders.

Inbreds from DEKALB and Holden's tend to be scattered across the dendrogram, although Holden's lines compose more than half of the Mo17 cluster (Fig. 1). Ten of the 13 Novartis lines cluster with Stiff Stalk germplasm;

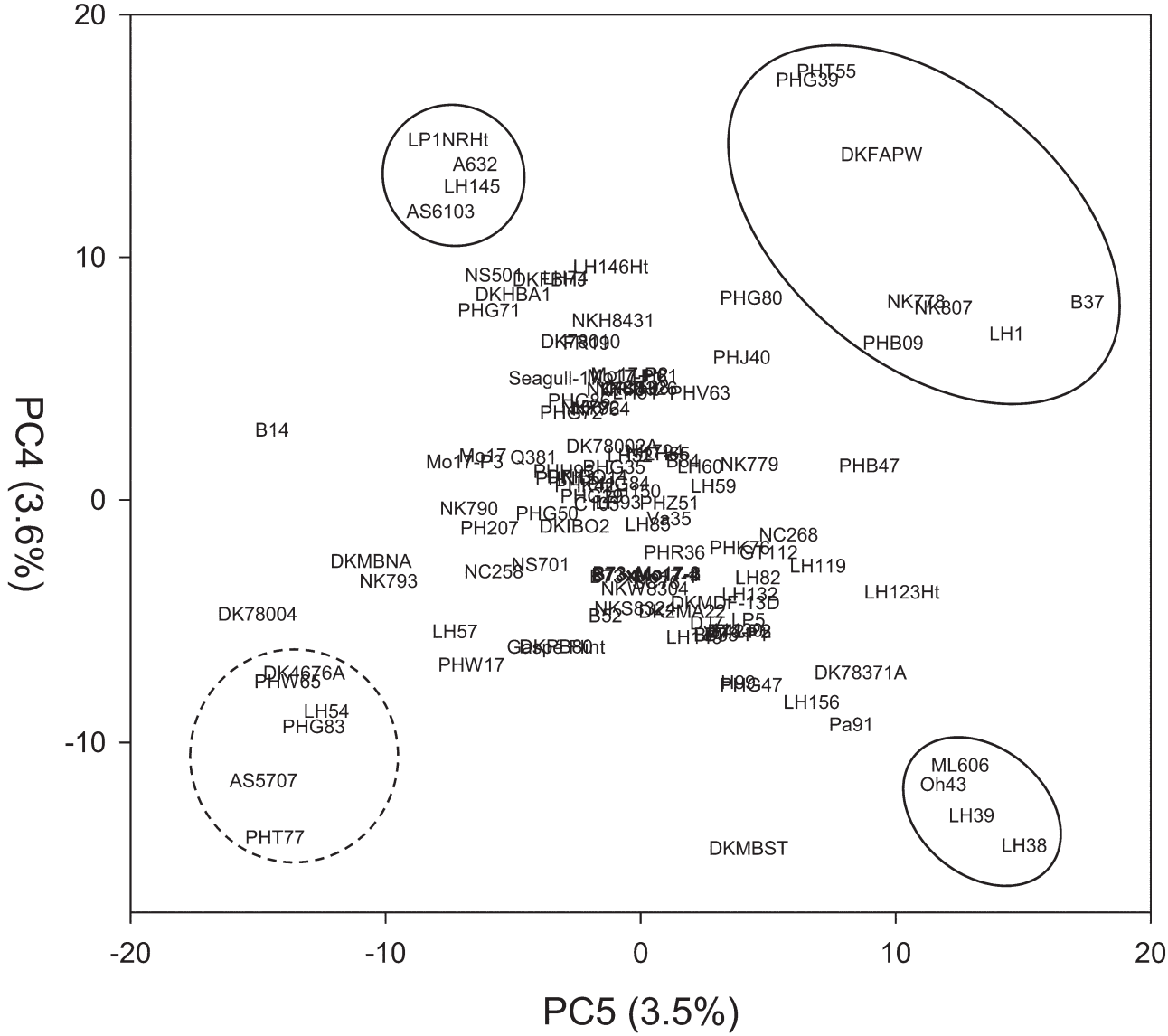

Figure 3. Scatter plot of the fourth and fifth principal components (PC4, PC5), which provides additional separation between B37, A632, and Oh43 clusters. A fourth cluster (dashed line) is evident; these lines are primarily of mixed or unrelated background.

the remaining three cluster among Lancaster lines. Pioneer lines have a strong presence in the Iodent and B37 clusters and are generally absent from the other clusters, with the exception of PHW17 and PHG86 in the B73 cluster and PHG47 and PHK76 near the Oh43 cluster. These observations are consistent with findings by Mikel and Dudley (2006). It is important to remember that inferences are restricted to this particular set of ex-PVPA inbreds and may not necessarily translate to the applicant company's germplasm base as a whole.

Most of the variation between duplicate B73 and Mo17 samples is attributed to variation between different seed sources of the same inbred, a phenomenon documented by Gethi et al. (2002). To a lesser degree, variation between B73 and Mo17 samples with the same seed source is attributed to marker-scoring error. For example, among the four Mo17 samples, the average Jaccard's similarity between the NCSU sample and the three Pioneer samples is 0.963 while the average similarity among the three replicate Pioneer samples is 0.997.

Principal Components Analysis

Principal components analysis yields separation of the lines into clusters similar to those in UPGMA cluster analysis. A three-dimensional scatter plot of the first three principal components (PC1–PC3) resembles a tetrahedron, the edges of which are formed by clusters that are represented by B73 (Stiff-Stalk), Mo17 (Lancaster), and PH207 (Iodent) (Fig. 2). The A632, Oh43, and B37 clusters are located near the vertex of the tetrahedron, though B37 and, to a greater extent, A632, are shifted toward the B73 cluster. Additional separation between the A632, Oh43, and B37 clusters is seen along the fourth and fifth principal components (PC4–PC5; Fig. 3).

The first five principal components explain 32.6% of the variation across the 614 SNPs included in PCA. The

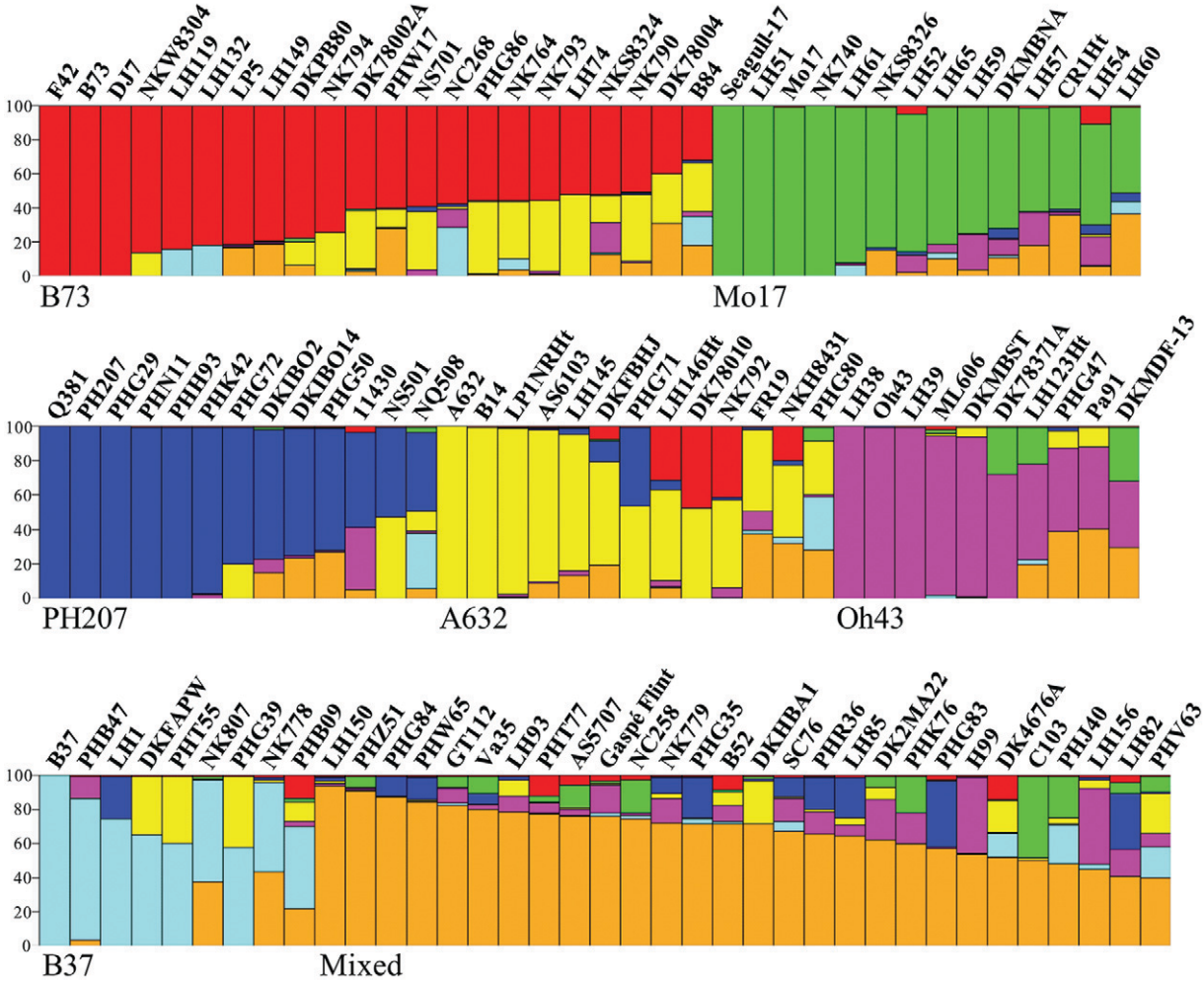

Figure 4. Bar plot of maize inbred membership in seven inferred populations based on admixture values, *Q*, from model-based clustering. Each bar represents an inbred, and each inbred's membership in a population is represented proportionally by color: B73 (red), Mo17 (green), PH207 (navy blue), A632 (yellow), Oh43 (fuchsia), B37 (light blue), and mixed (orange).

first principal component distinguishes between B73-type and Mo17-type lines. Along the PC1 axis, Mo17 lies at one extreme, B73 at the other extreme, and B73 × Mo17 halfway in between. PH207-type lines fall near 0.0 along PC1. The second principal component separates PH207-type lines from B73-type and Mo17-type lines with B73 × Mo17, B73, and Mo17 at one extreme and Q381 at the other extreme. The third principal component provides separation within the B73, Mo17, and Iodent clusters. The fourth principal component provides additional Stiff-Stalk vs. Lancaster separation as the Oh43 cluster is distinguished from the A632 and B37 clusters along this principal component. The fifth principal component separates the A632 cluster from the B37 cluster.

By definition, PC1 through PC5 explain progressively less of the variation across the entire marker data set. Thus, given our interpretation of these principal components, we can surmise that much of the variation across this set of ex-PVPA inbreds is attributed to the following four clusters or cluster contrasts, in decreasing order of magnitude: B73 vs. Mo17; PH207; A632 and B37 vs. Oh43; and A632 vs. B37.

Population Structure

Analysis of population structure reveals population groupings that agree with those found with UPGMA cluster analysis and PCA. In model-based clustering with the software Structure 2.1, optimal population structure based on maximum likelihood estimates is achieved when seven populations are assumed, represented by B73, Mo17, PH207, A632, Oh43, B37, and a mixed population. The lines are listed by their primary membership in each of these populations in Table 5. Where admixture values, *Q*, for the population of primary membership are <0.50, lines are considered mixed within their population of primary membership. Each line's estimated membership in each of the seven populations is given in a bar plot in Fig. 4 (see also Supplementary Table S5).

The proportional contribution of each of the seven populations to the set of ex-PVPA and public lines evaluated in this study is B73 (15.9%), A632 (14.8%), Mo17 (12.9%), PH207 (12.7%), Oh43 (11.2%), B37 (8.0%), and mixed (24.4%). The least significant difference estimate for comparison among these proportions is 3.0% ($\alpha = 0.05$). If

Table 3. Subsets of maize lines used in diversity analysis.

DEKALB Genetics	Holden's Foundation Seeds		Pioneer Hi-Bred Int.		Novartis Seeds	Other	Public
DK2MA22	LH1	LH119	PH207	PHJ40	NK740	11430	A632
DK4676A	LH38	LH123Ht	PHB09	PHK42	NK764	AS5707	B14
DK78002A	LH39	LH132	PHB47	PHK76	NK778	AS6103	B37
DK78004	LH51	LH145	PHG29	PHN11	NK779	CR1Ht	B52
DK78010	LH52	LH146Ht	PHG35	PHR36	NK790	DJ7	B73
DK78371A	LH54	LH149	PHG39	PHT55	NK792	F42	B84
DKFAPW	LH57	LH150	PHG47	PHT77	NK793	FR19	C103
DKFBHJ	LH59	LH156	PHG50	PHV63	NK794	LP1NRHt	H99
DKHBA1	LH60		PHG71	PHW17	NK807	LP5	Mo17
DKIBO2	LH61		PHG72	PHW65	NKH8431	ML606	NC268
DKIBO14	LH65		PHG80	PHZ51	NKS8324	NQ508	Oh43
DKMBNA	LH74		PHG83		NKS8326	NS501	Pa91
DKMBST	LH82		PHG84		NKW8304	NS701	Va35
DKMDF-13D	LH85		PHG86			Q381	
DKPB80	LH93		PHH93			Seagull-17	

Table 4. Public maize inbreds cited as prominent in seven papers from 1984 to 2006. Inbreds are only included if they were cited by more than one author.

Inbred	Baker (1984)	Darrah & Zuber (1986)†	Smith (1988)	Goodman (1992)	Troyer (1999)	Lu and Bernardo (2001)	Mikel and Dudley (2006)‡
A632	x	x	x	x	x		
A634	x	x					x
B14 (B14A)				x	x	x	x
B37		x		x	x	x	x
B73	x	x	x	x	x	x	x
B84				x		x	x
C103				x	x	x	
CM105	x						x
H99	x					x	x
Mo17	x	x	x	x	x	x	x
Oh43	x§		x	x	x	x	
W117	x						x

†Includes only the top five inbreds noted by Darrah and Zuber (1986).

‡Described by Mikel and Dudley (2006) as the most frequently used public inbreds as cited in Plant Variety Protection Act and U.S. Patent certificates.

§In an unpublished version of Baker's (1984) paper, LH38 (closely related to Oh43) was prominent.

the populations are consolidated into the three predominant heterotic groups represented in the United States, Stiff Stalk (B73, B37, and A632), Lancaster (Mo17 and Oh43), and Iodent (PH207), then the relative contributions are Stiff Stalk (38.7%), Lancaster (24.1%), and Iodent (12.7%); all proportions are significantly different from each other at $\alpha = 0.01$. There are several lines in Fig. 4 that have primary membership in the mixed population but have a large proportion of secondary membership in another population. Among these lines are the following, listed with the population of secondary membership: LH82 and PHG83 (PH207), C103 (Mo17), and H99 and LH156 (Oh43). These membership assignments are expected, given the pedigrees of these lines (see Table 1 in Mikel, 2006).

There is notable overlap between the B73 and A632 populations in Fig. 4. This is expected because the pedigree of A632 is 93.8% B14 [(Mt42 × B14)B14^3], and B14 and B73 are derived from the first and fifth cycles, respectively, of the Iowa Stiff Stalk Synthetic (Gerdes et al., 1993).

Line Derivation

Pedigrees for some of the ex-PVPA inbreds, as provided in PVPA certificates, are ambiguous, incomplete, erroneous, or include progenitors of an undisclosed origin. In these cases, results from marker-based clustering can provide additional insight and clarification about line derivation. For example, the pedigree of PHG39 (A33GB4 × A34CB4) is comprised of two proprietary Pioneer lines, previously derived from Stiff Stalk germplasm. In the current analysis, PHG39 clusters with B37-type lines and analysis with Structure 2.1 gives PHG39 a 99.7% Stiff Stalk membership (58.0% B37-type, 41.6% A632-type, and 0.1% B73-type). Similarly, Q381 is described in its PVPA certificate as an off-type of Pioneer hybrid 3369. However, in all analyses Q381 clusters with the Iodents, very close to PH207 (similarity coefficient of 0.98). Moreover, PH207 is not a parent of Pioneer hybrid 3369. It is reasonable to assume that while Q381 must have been retrieved from a Pioneer hybrid, given its similarity to PH207 (the likely female parent of said hybrid), it was not recovered from Pioneer hybrid 3369 as indicated in the PVPA record. Background clarification

for 19 ex-PVPA inbreds that have otherwise conflicting, ambiguous, or undisclosed pedigrees is provided in Table 6 (compare with Table 1 and Table 1 in Mikel, 2006).

In all three cluster analyses, there were certain lines that did not fit into any of the defined clusters or populations, falling between groups in UPGMA and PCA or within the mixed population in model-based clustering. In model-based clustering, the mixed population accounted for 24.4% of background assignment. Many of the ambiguities among these lines might be resolved if additional key progenitors are considered. However, given the lack of distinct clustering among these unassigned lines, either in UPGMA clustering, PCA, or model-based clustering, additional progenitors would likely provide clarification for small subsets of lines. The consideration of additional key progenitors might also aid in further dissecting line derivation. For instance, Smith et al. (1997) indicate that PHG39 has parentage from Maiz Amargo. However, since Maiz Amargo was not included in our genotyping, PHG39 retains a primary Stiff Stalk membership, as expected.

Diversity

Genetic diversity among the ex-PVPA lines as a whole was assessed via gene diversity (0.346), mean Jaccard's distance (0.514), and average allele number per locus (1.87) (Table 7). The values of these parameters are really only meaningful for comparisons among subsets. Within this set of ex-PVPA inbreds, genetic diversity was compared among subsets of lines, divided by application company: DEKALB, Holden's, Novartis, Pioneer, other, and public (Table 3). Comparisons are summarized in Table 7. For gene diversity and average Jaccard's distance, there are minimal differences between companies, though in both cases Novartis shows significantly less diversity than at least two other companies.

Allelic diversity among subsets of the ex-PVPA lines was quantified by measuring (i) average allele number per locus; (ii) unique alleles, alleles that are present in only one of the six subsets; (iii) absent alleles, alleles that are present in all but one of the subsets; and (iv) total absent alleles, the total number of alleles that are absent in a given subset but present in at least one other subset (Table 7). A resampling procedure was used to minimize sample-size bias in comparison between subsets. For average allele number, there are few statistically significant differences among subsets although Novartis has significantly fewer alleles per locus than the DEKALB and public subsets. For unique alleles, DEKALB has the highest average, 6.69. For absent alleles, the public subset has the fewest with an average of 5.99 and for total absent alleles DEKALB had has the fewest with an average of 55.22. When subsets are ranked by an index of proportional contribution to the total number of unique alleles, absent alleles, and total absent alleles, the DEKALB subset ranks the highest, followed by public, other, Pioneer, Holden's, and Novartis, in order of decreasing allelic diversity.

Table 5. Maize inbred membership in seven inferred populations from model-based clustering with Structure 2.1. Lines are grouped in their populations of primary membership. Where admixture values, *Q*, <0.50, lines are considered mixed within their population of primary membership.

Group	Lines
B73	B73, DJ7, DK78002A, DKPB80, F42, LH74, LH119, LH132, LH149, LP5, NC268, NK764, NK790, NK793, NK794, NKS8324, NKW8304, NS701, PHG86, PHW17
B73 Mixed	B84, DK78004
Mo17	CR1Ht, DKMBNA, LH51, LH52, LH54, LH57, LH59, LH60, LH61, LH65, Mo17, NK740, NKS8326, Seagull-17
PH207	11430, DKIBO2, DKIBO14, NS501, PH207, PHG29, PHG50, PHG72, PHH93, PHK42, PHN11, Q381
PH207 Mixed	NQ508
A632	A632, AS6103, B14, DK78010, DKFBHJ, LH145, LH146Ht, LP1NRHt, NK792, PHG71
A632 Mixed	FR19, NKH8431, PHG80
Oh43	DK78371A, DKMBST, LH38, LH39, LH123Ht, ML606, Oh43
Oh43 Mixed	PHG47, DKMDF-13D, Pa91
B37	B37, DKFAPW, LH1, NK778, NK807, PHB47, PHG39, PHT55
B37 Mixed	PHB09
Mixed	AS5707, B52, C103, DK2MA22, DK4676A, DKHBA1, Gaspé Flint, GT112, H99, LH82, LH85, LH93, LH150, LH156, NC258, NK779, PHG35, PHG83, PHG84, PHJ40, PHK76, PHR36, PHT77, PHV63, PHW65, PHZ51, SC76, Va35

Table 6. Backgrounds for ex-Plant Variety Protection Act maize inbreds with conflicting, ambiguous, or undisclosed pedigrees.

Inbred†	Marker-inferred background (proportion)‡						
	B73	Mo17	PH207	A632	Oh43	B37	Mixed
11430	0.03	0.01	0.55	0.00	0.36	0.00	0.05
DK4676A	0.14	0.01	0.00	0.19	0.01	0.14	0.52
DK78371A	0.00	0.27	0.00	0.00	0.72	0.00	0.00
DKIBO2	0.01	0.01	0.75	0.00	0.08	0.01	0.15
DKIBO14	0.00	0.01	0.74	0.01	0.01	0.00	0.24
DKMBNA	0.00	0.71	0.06	0.00	0.10	0.01	0.11
LH38	0.00	0.00	0.00	0.00	1.00	0.00	0.00
LH39	0.00	0.00	0.00	0.00	0.99	0.00	0.00
LH82	0.04	0.06	0.33	0.00	0.15	0.01	0.41
LH123Ht	0.00	0.21	0.00	0.01	0.55	0.03	0.21
NK792	0.41	0.00	0.01	0.51	0.06	0.00	0.00
NQ508	0.00	0.03	0.45	0.12	0.01	0.32	0.06
NS501	0.00	0.00	0.53	0.47	0.00	0.00	0.00
PHG39	0.00	0.00	0.00	0.42	0.00	0.58	0.00
PHJ40	0.01	0.24	0.00	0.04	0.00	0.23	0.49
PHK76	0.00	0.21	0.00	0.00	0.18	0.00	0.60
PHT55	0.00	0.00	0.00	0.39	0.00	0.60	0.00
PHV63	0.01	0.09	0.01	0.24	0.08	0.18	0.40
Q381	0.00	0.00	1.00	0.00	0.00	0.00	0.00

†A similar table for all 109 inbreds included in model-based clustering can be found in Supplementary Table S5.

‡Background proportions are based on admixture values, *Q*, from model-based clustering with Structure 2.1.

Table 7. Measures of genetic diversity among subsets of maize lines as defined in Table 3. With the exception of absent alleles and total absent alleles, higher values indicate greater diversity.

Company†	No. of lines‡	Avg. gene diversity	Avg. Jaccard's distance	Avg. allele no.	Unique alleles§	Absent alleles¶	Total absent alleles#
Comparison between six subsets, five ex-PVPA†† subsets and the public subset							
DEKALB	15	0.328	0.517	1.89	6.69	10.90	55.22
Holden's	23	0.330	0.510	1.85	1.00	10.00	77.70
Novartis	13	0.298‡‡	0.476§§	1.81¶¶	2.79	22.93	102.10
Pioneer	26	0.310	0.497	1.85	5.11	24.11	78.93
Other	15	0.323	0.506	1.86	1.51	8.87	70.70
Public	13	0.332	0.519	1.88	5.56	5.99	59.10
Comparison between the public subset and all ex-PVPA inbreds							
Public	13	0.332	0.519	1.88	40.20	31.47	31.47
All ex-PVPA	92	0.346	0.514	1.87	31.47	40.20	40.20

†DEKALB, DEKALB Genetics; Holden's, Holden's Foundation Seeds; Novartis, Novartis Seeds; Pioneer, Pioneer Hi-Bred Int.

‡Average allele no., unique alleles, absent alleles, and total absent alleles were calculated by resampling groups of 13 lines, 5000 times.

§Alleles that are present in only the given subset.

¶Alleles that are absent in the given subset but present in all other subsets.

#Total number of alleles that are absent in a given subset but present in at least one other subset.

††PVPA, Plant Variety Protection Act.

‡‡Novartis is significantly different from DEKALB at $\alpha = 0.05$ and Holden's and Public at $\alpha = 0.01$.

§§Novartis is significantly different from DEKALB at $\alpha = 0.05$.

¶¶Novartis is significantly different from DEKALB and Public at $\alpha = 0.01$.

When the public subset is compared against all of the ex-PVPA inbreds collectively, there are no significant differences in average gene diversity, average Jaccard's distance, or average allele number (Table 7). However, the public subset has more unique alleles and fewer absent alleles than the all ex-PVPA subset (for a two-way comparison, the number of total absent alleles, by the definition here, is equal to the number of absent alleles).

Among all six subsets there are 18 unique alleles, 71 absent alleles, and 377 total absent alleles. The mean allele frequency among the 18 unique alleles is 0.068, ranging from 0.022 to 0.154. These alleles are listed with positions and frequencies in Table 8; allelic frequencies corresponding to the set of temperate maize inbreds currently found in the Panzea database are also included.

There are 64 alleles present in the ex-PVPA germplasm that are absent in our public subset. When Gaspé Flint, GT112, NC258, and SC76 are added to the public subset, this number drops to 34 (Table 9). Each of these 34 alleles is present in the public temperate inbreds currently in the Panzea database, having a mean frequency of 0.080 and a range of 0.007 to 0.231. The mean frequency for these alleles is significantly greater (0.033, $P = 0.01$) in the temperate Panzea inbreds than in the ex-PVPA inbreds.

The ex-PVPA inbreds have little to offer in terms of allelic diversity for the 614 SNP markers used in this study. For each of the six measures of genetic diversity employed here, the public subset ranks first or second in diversity. Where alleles are unique to the ex-PVPA inbreds in our data set, the same alleles are found in broader sets of temperate public maize germplasm that are in the Panzea database. However, this may not mitigate the value of the ex-PVPA germplasm to public breeding programs because accumulation of favorable alleles, rather than single alleles per se, are generally of the greatest interest to breeders, especially when selecting for yield, grain moisture, and other quantitatively inherited traits.

There are limitations to the usefulness of the measures of genetic diversity employed in this study. Because genotyping was done with a bi-allelic assay, allele number is restricted to a maximum of two alleles per locus. Similarly, gene diversity, which is calculated on a per-locus basis, is constrained by allele number. Nevertheless, the large sample of SNP markers available for assay ameliorates these problems. The average Jaccard's distance is influenced by groups of closely related lines or the presence of a single diverse line. Investigation of genetic diversity based on haplotype structure may remedy some of the limitations of a bi-allelic assay, but was not addressed in this study.

There are also sampling limitations when measuring genetic diversity, and inferences are restricted to this particular set of ex-PVPA inbreds. First, lines evaluated in this study are dated; most entered PVPA protection between 1983 and 1989. Evidence presented by Duvick (1984) and Smith et al. (2004) reveals shifts in germplasm usage and diversity in time within breeding programs. Such shifts occurring within the past 20 yr would not be represented in these ex-PVPA inbreds. Second, the lines evaluated here are a small proportion of the approximately 800 inbreds in the PVPA system and are not necessarily a reflection of the germplasm that was available to private seed companies at the time these lines entered protection; many inbreds may

be protected via means other than PVPA such as trade secrets, license agreements, and patents. Some companies, such as Illinois Foundation Seeds (Champaign, IL) and Thurston Genetics (Olivia, MN), have essentially opted out of the PVPA program. Third, mergers and acquisitions within the maize seed industry in the last decade have drastically changed the maize breeding landscape. Some of the proprietary boundaries that once limited the flow of germplasm between breeding programs have now been dissolved. For example, when lines from Asgrow, DEKALB, and Holden's, all presently owned by Monsanto (Monsanto Company, St. Louis, MO), are considered collectively, Monsanto ranks second in allelic diversity, behind the public subset. Despite these sampling limitations, the ex-PVPA inbreds are the only publicly available representation of the germplasm being used on U.S. production acreage. Until regular assays of genetic diversity of U.S. maize hybrids can be conducted, as recommended by Mikel and Dudley (2006) and Smith (2007), the ex-PVPA inbreds, or inbreds that are currently under PVPA or U.S. Patent protection, are the best germplasm source for assessing genetic diversity in U.S. maize production.

Table 8. Genome position and frequency in subsets of ex–Plant Variety Protection Act lines for alleles that are unique to one of the six applicant company or public subsets. Allelic frequencies corresponding to the expanded set of temperate maize inbreds currently found in the Panzea database (http://www.panzea.org) are also given.

Locus	Nucleotide	Position†	Subset‡	Frequency	
				Subset	Temperate maize§
PHM4313-17	T	1-260.72	Public	0.077	0.102
PHM11114-25	T	1-476.82	Public	0.083	0.068
PZA00245-14	C	1-972.99	Pioneer	0.040	0.019
PHM482-23	G	2-122.41	DEKALB	0.036	0.080
PHM8352-4	G	2-342.55	Novartis	0.077	0.069
PZA03083-7	C	2-373.69	Pioneer	0.042	0.007
PZA00367-6	T	2-375.32	Pioneer	0.050	0.025
PHM8283-23	A	2-591.47	Public	0.091	0.186
PHM13942-10	A	3-346.79	DEKALB	0.143	0.068
PZA00220-11	C	3-597.59	Pioneer	0.040	0.054
PZA00289-11	T	4-300.73	Public	0.083	0.231
PZA00233-8	A	4-362.32	Pioneer	0.022	0.158
PZA00069-4	T	5-664.3	Pioneer	0.083	0.120
PZA00543-2	C	6-116.22	Novartis	0.077	0.014
PZA02970-9	A	6-165.58	Pioneer	0.043	0.117
PZA00533-3	C	7-330.61	DEKALB	0.077	0.057
PZA00255-15	T	Conflict¶	Pioneer	0.038	0.033
PZA00378-9	T	Conflict	DEKALB	0.045	0.147

†IBM2 chromosome-position (Lee et al., 2002).

‡DEKALB, DEKALB Genetics; Novartis, Novartis Seeds; Pioneer, Pioneer Hi-Bred Int.

§Temperate maize includes all public temperate inbreds in the Panzea marker database for the respective marker.

¶Conflicting positions given in Panzea marker database.

CONCLUSIONS AND INFERENCES

As increasingly more private research dollars are invested in maize breeding, public programs become more obsolete in terms of commercial maize production. Funding of public line development programs has essentially evaporated and few public programs remain in the United States. Consequently, there is less improved germplasm available for maize breeding at public institutions. As germplasm has become increasingly proprietary and as public inbreds receive less attention, the flow of germplasm into and between private breeding programs has been quelled. The formerly free exchange of germplasm among public maize breeding programs has also been hindered by the adoption of intellectual property protection policies at many public institutions.

The recent wave of PVPA certificate expirations represents a newly available germplasm resource for public breeding programs and for private breeding programs where previous access was restricted. The ex-PVPA lines may be valuable to individual breeding programs as a new source of genetic variation for traits of agronomic importance. However, the ex-PVPA lines are three to six cycles behind current elite germplasm in overall improvement and do not represent additional diversity for the U.S. maize germplasm base as a whole. How this germplasm is used in breeding efforts will depend on the breeding objectives and needs of individual programs and may differ between private and public breeding programs.

The role of public maize breeding programs in the modern maize breeding landscape is worth examination. The improvement gap between public and private germplasm continues to widen and the use of 20-yr-old ex-PVPA inbreds in public programs will not reverse the trend. Further, individual commercial programs will presumably rework ex-PVPA inbreds, particularly where they are from a genetic background that is unique to the program of interest. One should consider whether public breeding programs can most beneficially contribute through recycling favorable linkage blocks present in these ex-PVPA materials or through considering more exotic, less adapted sources of germplasm that contain diverse alleles not yet leveraged in U.S. commercial materials. Maize breeders have long advocated the use of exotic germplasm in temperate breeding efforts (Melhus, 1948), but relatively little has been done in this regard (Goodman, 1998). One exception is the Germplasm Enhancement of Maize (GEM) project, from which several dozen lines of 25 or 50% tropical origin have been released (Abel et al., 2001; Hawk and Weldekidan, 2005; Pratt et al., 2005; Balint-Kurti et al., 2006; Carson et al., 2006; Campbell et

Table 9. Genome position and allele frequency in the entire set of ex–Plant Variety Protection Act (PVPA) lines for alleles that are absent in the public maize inbreds but are present in the ex-PVPA inbreds. Allelic frequencies corresponding to the expanded set of temperate maize inbreds currently found in the Panzea database (http://www.panzea.org) are also given.

Locus	Nucleotide	Position†	Frequency	
			Ex-PVPA	Temperate maize‡
PZA02921-9	T	1-170.03	0.071	0.128
PZA00714-1	T	1-392.75	0.158	0.231
PHM4185-17	T	1-405.02	0.172	0.106
PZA03028-5	A	1-450.84	0.115	0.055
PHM482-23	G	2-122.41	0.006	0.080
PHM4780-38	G	2-251.13	0.048	0.074
PHM3931-17	C	2-342.45	0.033	0.088
PHM883-16	A	2-345.29	0.016	0.107
PZA02982-5	T	2-346.4	0.021	0.038
PZA03083-7	C	2-373.69	0.011	0.007
PZA00367-6	T	2-375.32	0.014	0.025
PHM13942-10	A	3-346.79	0.023	0.068
PZA02952-10	T	3-540.2	0.018	0.024
PHM3352-19	A	3-617.49	0.028	0.061
PZA00233-8	A	4-362.32	0.006	0.158
PHM3590-19	A	4-373.45	0.024	0.071
PHM5780-13	T	4-619.4	0.042	0.051
PZA00543-2	C	6-116.22	0.011	0.014
PZA02970-9	A	6-165.58	0.012	0.117
PHM13451-15	C	6-235.82	0.061	0.056
PZA00533-3	C	7-330.61	0.011	0.057
PZA00655-1	T	7-347.18	0.017	0.127
PHM2691-31	A	7-86.33	0.172	0.071
PHM3856-10	G	8-274.42	0.017	0.148
PZA02811-4	A	8-316.24	0.028	0.103
PZA00686-4	C	8-388.8	0.030	0.019
PZA00310-5	T	10-173.54	0.030	0.117
PZA00587-3	G	10-225.69	0.022	0.059
PZA00103-20	C	Conflict§	0.056	0.146
PZA00255-15	T	Conflict	0.011	0.033
PZA00615-3	A	Conflict	0.196	0.053
PHM1960-37	T	Unmapped	0.056	0.105
PHM6608-5	A	Unmapped	0.034	0.085
PZA03063-18	C	Unmapped	0.031	0.047

†IBM2 chromosome-position (Lee et al., 2002).

‡Temperate maize includes all temperate inbreds in the Panzea marker database for the respective marker.

§Conflicting positions given in Panzea marker database.

al., 2007). Similarly, the North Carolina State University maize breeding program has been working extensively with tropical germplasm for >30 yr and has released 90 inbreds of partial or all-tropical origin (Goodman, 1992; Nelson and Goodman, 2008). Perhaps public maize breeding programs should look 20 to 30 yr ahead and strategically use specific ex-PVPA lines as an adapted source in which to place useful exotic germplasm.

Effective utilization of the ex-PVPA maize inbreds in breeding will require accurate information on line derivation and performance. Inbred per se and topcross hybrid performance evaluations are most effectively done on a regional basis due to genotype × environment interactions. All of the ex-PVPA lines presented in this paper are currently in yield trials at North Carolina State University and multiyear data will be available by fall 2008. Pedigrees, as found in PVPA certificates, can be used to infer line derivation and relationships but are limited in their own right. Molecular markers are another valuable tool for inferring population structure and relationship among individual lines.

We have used molecular markers to evaluate the relationship and population structure across 92 ex-PVPA maize inbred lines. We have related these lines to well-known public inbreds and classified them relative to the following key progenitors: B73, Mo17, PH207, A632, Oh43, and B37. Several dozen ex-PVPA lines are of mixed or unrelated origin and do not fit into any one of the aforementioned groups. The inclusion of additional key progenitors in future research may aid in further dissecting the relationships among these lines. Our results support previous conclusions in the literature about the derivation of temperate U.S. maize germplasm (Table 4). We have also used results from the various cluster analyses to clarify ambiguities in pedigree records for some of the ex-PVPA inbreds (Table 6). While public germplasm does not lag behind the ex-PVPA germplasm in terms of genetic and allelic diversity, the ex-PVPA lines do represent a newly available source of elite maize germplasm that can likely add favorable variation for agronomic traits of interest to public and private maize breeding programs.

Acknowledgments

This research was supported by a USDA-IFAFS multidisciplinary training grant to North Carolina State University (award no. 2001-52101-11507), a National Science Foundation grant (DBI-0321467), the USDA-ARS GEM Project (SCA 3625-4-125), a Pioneer Fellowship in Plant Sciences (to P.T. Nelson), and an NC Ag Foundation Fellowship (to N.D. Coles). Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University, which is partially funded by Microsoft Corporation. Genotyping was conducted by Pioneer Hi-Bred International Inc. Some of the data used in this study were generated by the Panzea project (http://www.panzea.org, verified 8 Apr. 2008).

References

Abel, C.A., L.M. Pollak, W. Salhuana, M.P. Widrlechner, and R.C. Wilson. 2001. Registration of GEMS-0001 maize germplasm resistant to leaf blade, leaf sheath, and collar feeding by European corn borer. Crop Sci. 41:1651–1652.

Baker, R. 1984. Some of the open pollinated varieties that contributed the most to modern hybrid corn. Illinois Corn Breeder's School Proc. 20:1–19.

Balint-Kurti, P.J., M. Blanco, M. Millard, S. Duvick, J.B. Holland, M.J. Clements, R.N. Holley, M.L. Carson, and M.M. Goodman. 2006. Registration of 20 GEM maize breeding germplasm lines adapted to the southern USA. Crop Sci. 46:996–998.

Bernardo, R. 1993. Estimation of coefficient of coancestry using molecular markers in maize. Theor. Appl. Genet. 85:1055–1062.

Campbell, M.R., J. Jane, L. Pollak, M. Blanco, and A. O'Brien. 2007. Registration of maize germplasm line GEMS-0067. J. Plant Registrations 1:60–61.

Carson, M.L., P.J. Balint-Kurti, M. Blanco, M. Millard, S. Duvick, R.N. Holley, J. Hudyncia, and M.M. Goodman. 2006. Registration of nine high-yielding tropical by temperate maize germplasm lines adapted for the southern USA. Crop Sci. 46:1825–1826.

Coors, J.G. 2006. Who are plant breeders, what do they do, and why? p. 51–60. *In* K.R. Lamkey and M. Lee (ed.) Plant breeding: The Arnel R. Hallauer international symposium. Blackwell, Ames, IA.

Darrah, L.L., and M.S. Zuber. 1986. 1985 United States farm maize germplasm base and commercial breeding strategies. Crop Sci. 26:1109–1113.

Duvick, D.N. 1984. Genetic diversity in major farm crops on the farm and in reserve. Econ. Bot. 38:161–178.

Fan, J.B., A. Oliphant, R. Shen, B.G. Kermani, F. Garcia, K.L. Gunderson, M. Hansen, F. Steemers, S.L. Butler, P. Deloukas, L. Galver, S. Hunt, C. McBride, M. Bibikova, T. Rubano, J. Chen, E. Wickham, D. Douget, W. Chang, D. Campbell, B. Zhang, S. Kruglyak, D. Bentley, J. Haas, P. Rigault, and L. Zhou. 2003. Highly parallel SNP genotyping. Cold Spring Harbor Symp. Quant. Biol. 68:69–78.

Gerdes, J.T., C.F. Behr, J.G. Coors, and W.F. Tracy. 1993. Compilation of North American maize breeding germplasm. CSSA, Madison, WI.

Gethi, J.G., J.A. Labate, K.R. Lamkey, M.E. Smith, and S. Kresovich. 2002. SSR variation in important U.S. maize inbred lines. Crop Sci. 42:951–957.

Goodman, M.M. 1992. Choosing and using tropical corn germplasm. Corn Sorghum Res. Conf. Proc. 47:47–64.

Goodman, M.M. 1998. Research policies thwart potential payoff of exotic germplasm. Diversity 14:30–35.

Hawk, J.A., and T. Weldekidan. 2005. Registration of DE3, DE4, and DE5 parental inbred lines of maize. Crop Sci. 45:2669–2670.

Lee, M., N. Sharopova, W.E. Beavis, D. Grant, M. Katt, D. Blair, and A. Hallauer. 2002. Expanding the genetic map of maize with the B73 × Mo17 (IBM) population. Plant Mol. Biol. 48:453–461.

Liu, K., and S.V. Muse. 2005. PowerMarker: Integrated analysis environment for genetic marker data. Bioinformatics 21:2128–2129.

Lu, H., and R. Bernardo. 2001. Molecular marker diversity among current and historical maize inbreds. Theor. Appl. Genet. 103:613–617.

McMaster, G.S., and W. Wilhelm. 1997. Growing degree-days: One equation, two interpretations. Agric. For. Meteorol. 87:291–300.

Melhus, L.E. 1948. Exploring the maize germ plasm of the tropics. Corn Sorghum Res. Conf. Proc. 3:7–19.

Mikel, M.A. 2006. Availability and analysis of proprietary dent corn inbred lines with expired U.S. plant variety protection. Crop Sci. 46:2555–2560.

Mikel, M.A., and J.W. Dudley. 2006. Evolution of North American dent corn from public to proprietary germplasm. Crop Sci. 46:1193–1205.

Nei, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.

Nelson, P.T., and M.M. Goodman. 2008. Evaluation of elite exotic maize inbreds for use in temperate breeding. Crop Sci. 48:85–92.

Pratt, R.C., L.M. Pollak, and K.T. Montgomery. 2005. Registration of maize germplasm line GEMS-0002. Crop Sci. 45:2130–2131.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Rohlf, F.J. 2007. NTSYSpc: Numerical taxonomy system. Version 2.20N. Exeter Publishing, Setauket, NY.

SAS Institute. 2003. SAS/Stat user's guide. Version 9.1.3. SAS Inst., Cary, NC.

Smith, J.S.C. 1988. Diversity of United States hybrid maize germplasm: Isozymic and chromatographic evidence. Crop Sci. 28:63–69.

Smith, J.S.C. 2007. Pedigree background changes in U.S. hybrid maize between 1980 and 2004. Crop Sci. 47:1914–1926.

Smith, J.S.C., E.C.L. Chin, H. Shu, O.S. Smith, S.J. Wall, M.L. Senior, S.E. Mitchell, S. Kresovich, and J. Ziegle. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparisons with data from RFLPS and pedigree. Theor. Appl. Genet. 95:163–173.

Smith, J.S.C., D.N. Duvick, O.S. Smith, M. Cooper, and L. Feng. 2004. Changes in pedigree backgrounds of Pioneer brand maize hybrids widely grown from 1930 to 1999. Crop Sci. 44:1935–1946.

Sprague, G.F. 1971. Genetic vulnerability in corn and sorghum. Corn and Sorghum Res. Conf. Proc. 26:96–104.

Steel, R.G.D., J.H. Torrie, and D.A. Dickey. 1997. Principles and procedures of statistics: A biometrical approach. 3rd ed. McGraw-Hill, Boston, MA.

Troyer, F.A. 1999. Background of U.S. hybrid corn. Crop Sci. 39:601–626.

Zuber, M.S. 1975. Corn germplasm base in the U.S.: Is it narrowing, widening, or static? Corn and Sorghum Res. Conf. Proc. 30:277–286.

Zuber, M.S., and L.L. Darrah. 1980. 1979 U.S. corn germplasm base. Corn and Sorghum Res. Conf. Proc. 35:234–249.